(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
3 May 2001 (03.05.2001)

PCT

(10) International Publication Number
WO 01/31906 A2

(51) International Patent Classification[7]: H04N

(21) International Application Number: PCT/US00/41416

(22) International Filing Date: 23 October 2000 (23.10.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/161,149 23 October 1999 (23.10.1999) US

(71) Applicant: FASTVIDEO LLC [US/US]; 5565 Sterrett Place, Suite 322, Columbia, MD 21044 (US).

(72) Inventors: TOPIWALA, Pankaj, N.; 6429 Prestwick Drive, Clarksville, MD 21029 (US). TRAN, Trac, D.; 12225 Little Patuxent Parkway #L, Columbia, MD 21044 (US).

(74) Agent: WILLIAMS, Frederick, C.; Williams & Associates, Suite 300, 1030 Fifteenth Street, N.W., Washington, DC 20005-1501 (US).

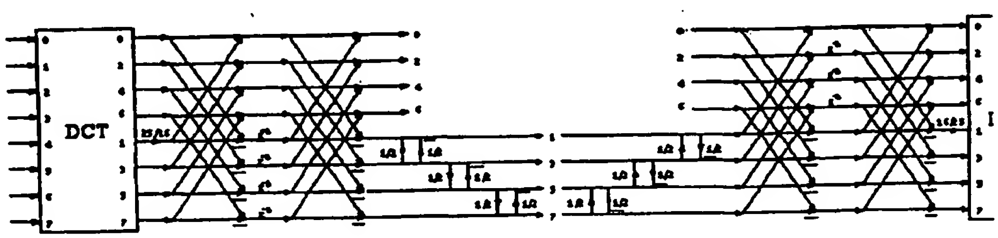
(81) Designated States (*national*): AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *Without international search report and to be republished upon receipt of that report.*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: METHOD FOR LOCAL ZEROTREE IMAGE CODING



(57) Abstract: A transform-based image compression framework called local zerotree (LZT) coding partitions the transform coefficients into small groups, each of which is encoded independently using popular zerotree algorithms. The LZT coder achieves similar coding performances as current state-of-the-art embedded coders. The advantage of this coding method is fourfold: (i) because of the reduction of memory buffering, LZT reduces the complexity of the codec implementation and increases the speed of tHE zerotree algorithm significantly, especially in hardware; (ii) LZT processes large images under limited memory constraint; (iii) LZT supports parallel processing mode as long as the transform in use has that capability; and (iv) LZT facilitates the coding/decoding of regions of interest. The penalty in coding performance is minute compared to global zerotree predecessors in that there are only a few extra bytes of side information. This new coder provides numerous other advantages: faster and simpler hardware implementation, supporting parallel processing mode, facilitating region-of-interest coding/decoding, and capable of processing large images under limited memory constraint. The only properties that LZT sacrifices are full embededness and progressive transmission.

WO 01/31906 A2

METHOD FOR LOCAL ZEROTREE IMAGE CODING

PRIORITY

This application claims the benefit of priority from copending United States provisional application 60/161,149, filed October 23, 1999.

FIELD OF THE INVENTION

This invention relates to the processing of images such as photographs, drawings, and other two dimensional displays. It further relates to the processing of such images after they have been captured in digital format or after they have been converted to or otherwise expressed in digital format. This invention further relates to use of novel coding methods to increase the speed and compression ratio for digital image storage and transmission.

BACKGROUND OF THE INVENTION

Recently subband coding has emerged as the leading standardization candidate in future image compression systems due to the development of the discrete wavelet transform (DWT). The multiresolution characteristics of the wavelet transform establish an intuitive foundation on which simple yet sophisticated methods of encoding the transform coefficients are developed. Exploiting the relationship between the parent and the offspring coefficients in wavelet tree, zerotree-based wavelet coders, for example, J. M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Trans. on SP*, vol. 41, pp. 3445-3462, Dec. 1993 [hereinafter **EZW**]; A. Said and W. A. Pearlman, "A new fast and efficient image codec on set partitioning in hierarchical trees," *IEEE Trans on Circuits Syst. Video Tech.*, vol. 6, pp. 243-550, June 1996 [hereinafter **SPIHT**]; and "Compression with reversible embedded wavelets," RICOH Company Ltd. submission to ISO/IEC JTC1/SC29/WG1 for the JTC1.29.12 work item, 1995. http://www.crc.ricoh.com/CREW [hereinafter **CREW**] can effectively order the coefficients by bit planes and transmit more significant bits first. This elegant coding scheme results in an embedded bit stream (i.e., it can be truncated at any point by the decoder to yield the best corresponding reconstructed image) and many other interesting features such as exact bit rate control and idempotency capability. It became apparent that the same idea can be applied to block transform coefficients as well if these are grouped in a wavelet-like octave-band

fashion, e.g., Z. Xiong, O. Guleryuz, and M. T. Orchard, "A DCT-based embedded image coder," *IEEE SP Letters*, vol. 3, pp. 289-590, Nov. 1996; H. Malvar, "Lapped biorthogonal transforms for transform coding with reduced blocking and ringing artifacts," *ICASSP97*, Munich, April 1997;] T. Klausutis and V. Madisetti, "Adaptive lapped transform-based image coding," *IEEE SP Letters*, vol. 4, pp. 245-547, Sept. 1997; T. D. Tran and T. Q. Nguyen, "A lapped transform embedded image coder," *ISCAS*, Monterey, May 1998, as illustrated in Figure 1. Global information is fully taken into account in these subband coders, leading to excellent performance.

However, perfect embeddedness is achieved at a high computational cost. Such embeddedness requires a full buffering of all the transform coefficients. First of all, this requirement becomes burdensome when the input images have high resolution (in desktop publishing applications for example) and/or the decoder has limited resources. Second, a full buffering with multiple passes is very costly in hardware implementation. Finally, full buffering prevents the application of powerful parallel processing techniques.

It is therefore an object of the current invention to sacrifice progressive transmission capability for a faster coder with a lower implementation complexity and a comparable quality performance. It is a further object of the current invention to construct such a coder by partitioning the transform coefficients into small local groups and encoding them independently using popular zerotree-based algorithms. It is yet a further object of this invention to provide a coder with negligible performance penalty. It is another object of this invention to provide a coder with performance almost identical to the fully embedded version when the partition size is reasonably large. It is another object of this invention to maintain embeddedness within the local groups while losing global embeddedness in the final output bitstream.

SUMMARY OF THE INVENTION

The LZT coding algorithm of this invention has at least the following fundamental steps:

1) Partitioning the input image into many local blocks. (The size of the partition is chosen based on a tradeoff between the desired level of resolution and the availability of computing resources, which determine an allowable level of implementation complexity.)

2) Transforming each partition, that is, each local block, independently using a block coder. (Some special attention is required at the partition boundaries to ensure continuity in the reconstructed image.)

3) Quantizing the coefficients in each partition by the appropriate JND thresholds or by truncating a certain number of least significant bit planes.

4) Arranging the quantized coefficients in wavelet-like quad-tree fashion as shown in Figure 2 and encode each group separately using any zerotree coding algorithm.

5) Concatenating the bitstreams to yield the final output.

Most generally, this invention comprises the steps of partitioning the image into a plurality of local blocks, performing a block coding transform on each local block of the plurality of local blocks so as to produce a corresponding plurality of sets of local block transform coefficients using a block coding algorithm, quantizing the local block transform coefficients in each set of local block transform coefficients of the plurality of local block transform coefficients, arranging the plurality of sets of quantized local block transform coefficients in a wavelet-like quad-tree structure, and concatenating the bit-streams of the encoded quantized coefficients from each of the plurality of local blocks.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows the wavelet and block transform analogy.

Figure 2 is a diagram of LZT coding.

Figure 3 illustrates the LiftLT analysis/synthesis lattice.

Figure 4 shows reconstructed 512 X 512 Barbara images from coding examples compiled in Table 1.

DESCRIPTION OF THE PREFERRED EMBODIMENT

Embedded coding algorithms rely on the hierarchical tree structure of coefficients called a *wavelet tree*, defined as a set of wavelet coefficients from different scales that belong in the same spatial locality as demonstrated in Figure 1(a), where the tree in the horizontal direction **5** is circled. All of the coefficients in the lowest frequency band make up the *DC band* or the *reference signal* (located at the upper left corner) **6**. Besides these DC coefficients, in a wavelet tree of a particular direction, each lower-frequency *parent node* has four corresponding higher-frequency *offspring nodes* 7. All coefficients below a parent node in the same spatial locality is defined as its *descendants* **8**. In this invention a coefficient c_{ij} is defined to be *significant* with respect to a given threshold T if $c_{ij} \geq T$, and *insignificant* otherwise. Review of image statistics has shown that if a coefficient is insignificant, it is very likely that its offspring and descendents are insignificant as well. Exploiting this fact, sophisticated embedded zerotree coders can output a single binary marker to represent very efficiently a large, smooth image area (an insignificant tree). The details of zerotree algorithms are known in the art and can be found in such publications as EZW, SPIHT, and CREW.

The key to the high efficiency of zerotree coders is accuracy in predicting the correlation between coefficients in the same spatial locality. Observation shows minimal correlation between far-apart coefficient trees. Thus a zerotree algorithm can capture very high coding performance when local groups of coefficient trees are encoded independently.

Classic bit allocation must also be taken into account in local zerotree coding. Alternate solutions are available to distribute the bit budget appropriately between the local groups of coefficients. On the one hand, embedded zerotree coders can employ a very simple quantization scheme -- all subbands are scalar-quantized by a single step-size. Hence, prior to the encoding step, the same number of least significant bit planes of all coefficients from all LZT groups are truncated. On the other hand, the method can diverge from bit-rate control to quality control in a fashion similar to that of JPEG, W. B. Pennebaker and J. L. Mitchell, JPEG: Still Image Compression Standard, Van Nostrand Reinhold, 1993, where each subband is quantized by a predetermined step-size governed by the response of the human visual system.

The basic LZT structure is shown in Figure 2. The LZT partitions can be of different shapes and sizes. For example, the image can be encoded in strips. The output bitstream for each LZT partition must include the maximum magnitude of the partition's

coefficients (to be used as the initial threshold value by the decoder) and the mean of its DC subband. Furthermore, there must also be a special END-OF-BLOCK marker to inform the decoder of the boundary between different LZTs. Encoding these values takes a small number of extra bytes. The side information needed for LZT coding is negligible unless the LZT block size becomes small. Another interesting feature that LZT can support is region-of-interest coding/decoding: the LZT partitions can be quantized with different step sizes resulting in different levels of reproducibility.

Since LZT is primarily a block coding algorithm, the ideal candidates at the transform stage are block transforms. The DCT is inappropriate since its non-overlapping basis creates blocking artifacts between 8 x 8 coefficient blocks as well as between larger LZT partitions. In the current preferred embodiment of this invention, we employ a high-performance 8 x 16 lapped transform with integer coefficients called the LiftLT, T. D. Tran, "The LiftLT: fast lapped transforms via lifting steps," submitted to ICASSP, Phoenix, March 1999, whose fast and efficient implementation is depicted in Figure 3. This transform is also described in a prior patent application by the current inventors, Tran, Trac D. and Topiwala, Pankaj, "Fast Lapped Image Transforms Using Lifting Steps," U. S. Patent Application No. 09/212,210, filed December 12, 1998, which is incorporated by reference as though fully set forth herein. The small 8-pixel overlap facilitates the manipulation of LZT boundaries. The well known discrete wavelet transform can be used as well. However, to minimize the LZT's buffering size, this invention must use efficient block-based implementations such as C. Chrysafis and A. Ortega, "Line based reduced memory wavelet image compression," *Proc. DCC98*, Snowbird, UT, March 1998; P. Cosman and K. Zeger, "Memory constrained wavelet-based image coding," *IEEE SP Letters*, vol. 5, Sep. 1998.

Table 1 sets forth numerical comparison of the coder of this invention with fully embedded coders. For fair comparison, the same SPIHT quantizer and entropy coder are utilized to transfer the coefficients in all coding schemes. The coders in comparison (with the associated transforms) are:

1. 64 X 64 LZT coder, 8 X 16 LiftLT
2. 128 X 128 LZT coder, 8 X 16 LiftLT
3. 256 X 256 LZT coder, 8 X 16 LiftLT
4. fully embedded LT coder, 8 X 16 LiftLT
5. 64 X 64 LZT coder, 9/7-tap wavelet

6. fully embedded SPIHT coder, 9/7-tap wavelet.

TABLE 1

	Coding Algorithm	LZT	LZT	LZT	EZ-LT	LZT	SPIHT
	Local Block Size	64 x 64	128 x 128	256 x 256	--	64 x64	--
	Transform	8 x 16 LT	8 x 16 LT	8 x 16 LT	8 x 16 LT	9/7 WL	9/7 WL
Barbara	Output Bytes	7869	6908	6623	6601	5847	5177
	Comp. Ratio	33.32:1	37.95:1	39.58:1	39.71:1	44.83:1	50.64:1
	PSNR	28.11	28.07	28.06	28.06	25.87	25.87
Goldhill	Output Bytes	10806	9939	9558	9395	8065	7209
	Comp. Ratio	24.26:1	26.38:1	27.43:1	27.90:1	32.50:1	36.36:1
	PSNR	31.17	31.15	31.15	31.15	30.14	30.17
Lena	Output Bytes	8599	7688	7352	7281	6428	5570
	Comp. Ratio	30.49:1	34.10:1	35.66:1	36.00:1	40.78:1	47.06:1
	PSNR	33.49	33.46	33.43	33.47	32.46	32.46

In the lapped transform cases, the modified zerotree structure was used in which each block of transform coefficients were treated analogously to a full wavelet tree. The objective coding results (PSNR in dB) for standard 512 X 512 Lena, Goldhill, and Barbara test images are tabulated in Table 1. All computed PSNR quotes are obtained from a real compressed bit stream with all overheads included. As expected, the LZT coding performance decreases as the local block size becomes smaller. For LZT of reasonable size ($^{1}/_{16}$ th of the original image size and up), the amount of side information is negligible. Several reconstructed Barbara images using LZTs of different size and the 8 x 16 LiftLT are shown in Figure 5. LZT images are comparable in visual quality to the fully embedded EZ-LT coder image. Top left: 64 X 64 LZT **10**. Top right: 128 X 128 LZT **11**. Bottom left: 256 X 256 LZT **12**. Bottom right: fully embedded EZ-LT coder **13**. The boundaries between LZT partitions are imperceptible.

CLAIMS

We claim:

1. A method of coding a digital image to produce data compression, comprising the steps of

    a. Partitioning the image into a plurality of local blocks;

    b. Performing a block coding transform on each local block of the plurality of local blocks so as to produce a corresponding plurality of sets of local block transform coefficients using a block coding algorithm;

    c. Quantizing the local block transform coefficients in each set of local block transform coefficients of the plurality of local block transform coefficients;

    d. Arranging the plurality of sets of quantized local block transform coefficients in a wavelet-like quad-tree structure; and

    e. Concatenating the bit-streams of the encoded quantized coefficients from each of the plurality of local blocks.

2. The method of Claim 1 in which the partitioning is performed by means of a zerotree coding algorithm.

3. The method of Claim 1 in which the block coding transform is a LiftLT transform.

4. The method of Claim 1 in which the transform is chosen from the group: discrete wavelet transform; line based reduced memory wavelet image transform; memory constrained wavelet-based image coder; lapped biorthogonal transform.

5. The method of Claim 2 in which the zerotree coding algorithms is a bitplane embedded encoding system.

6. The method of Claim 2 in which the zerotree coding algorithm is chosen from the group: Shapiro's EZW zerotree algorithm, Said-Pearlman's SPIHT encoding algorithm, and CREW's zerotree algorithm.

1/4

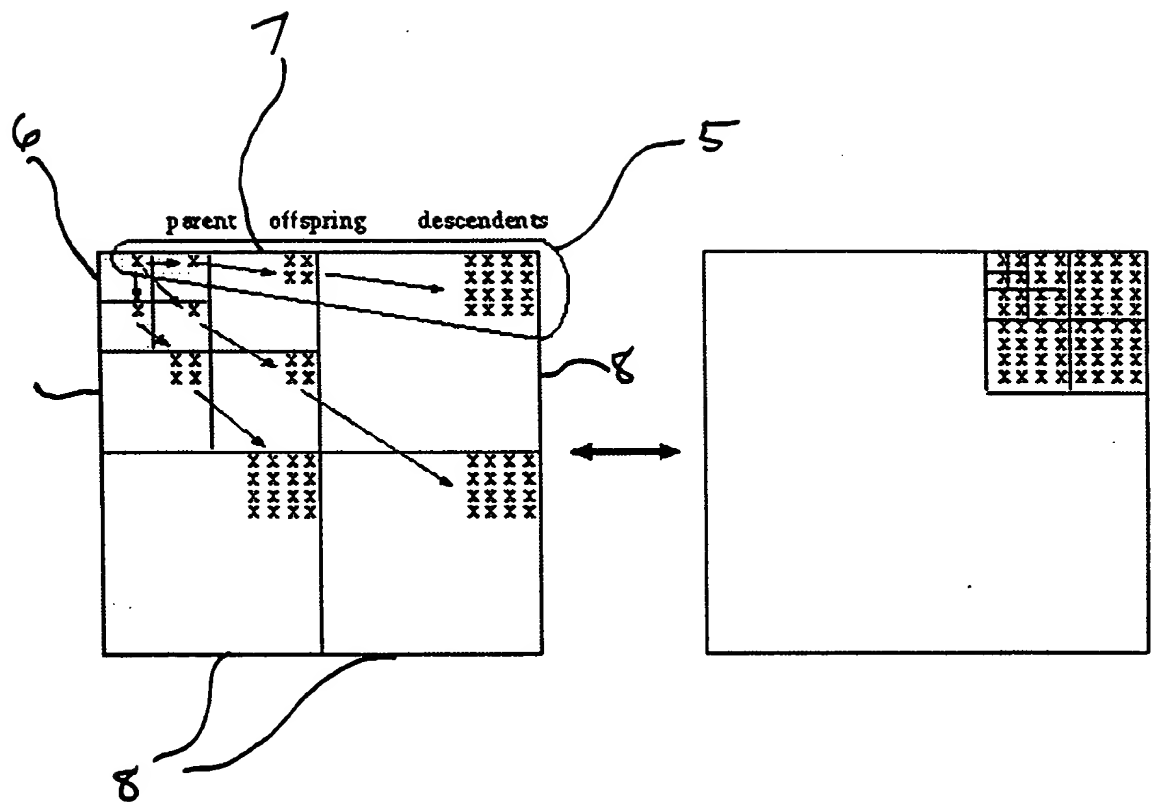


FIGURE 1

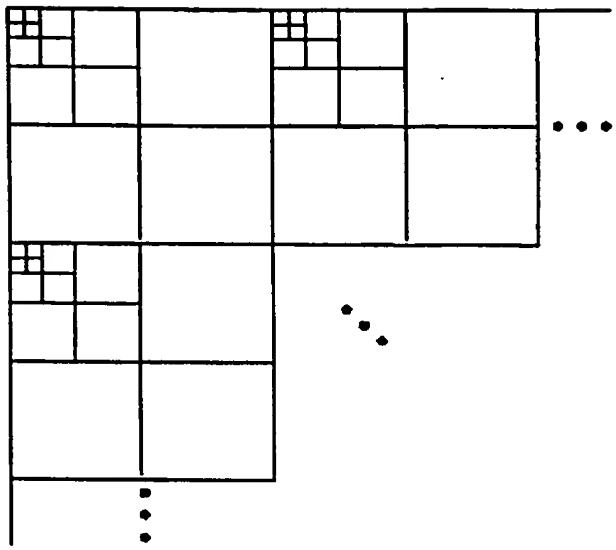


FIG. 2

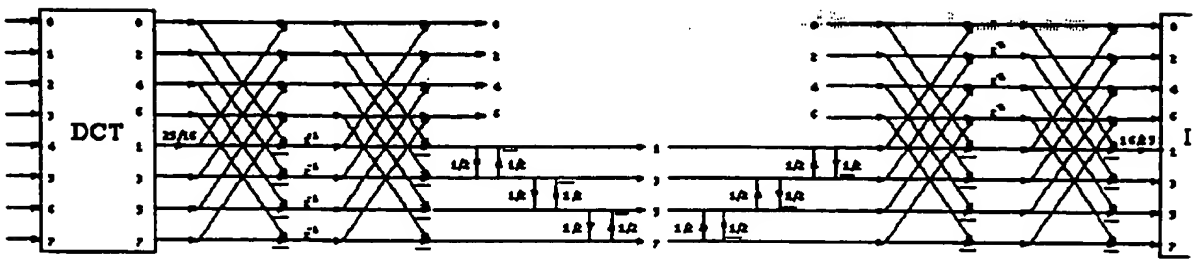


FIG. 3

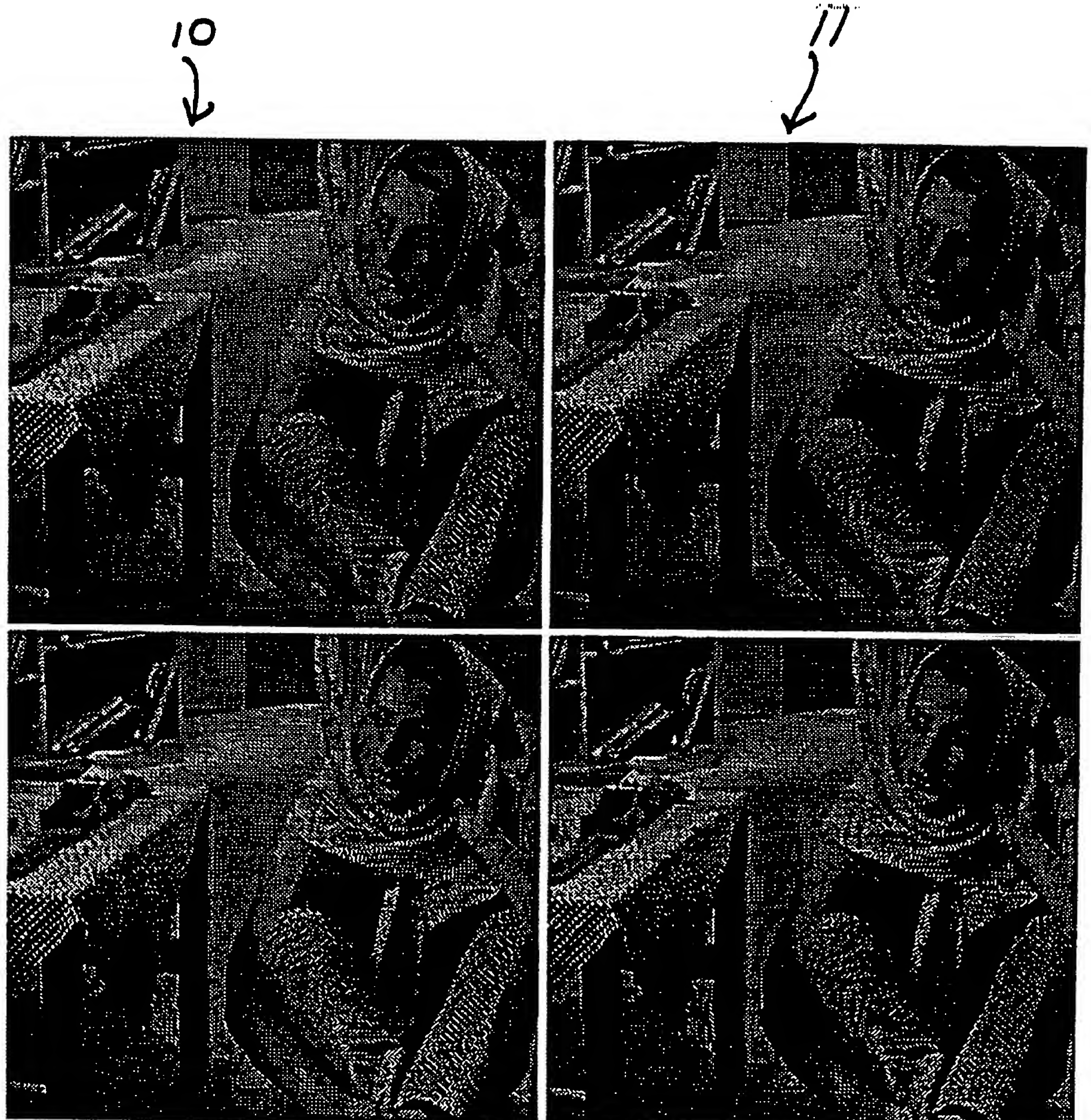


Figure 4. Reconstructed 512 × 512 Barbara images from coding example compiled in Table 1. Top left: 64 × 64 LZT. Top right: 128 × 128 LZT. Bottom left: 256 × 256 LZT. Bottom right: fully embedded EZ-LT coder.

12

13

FIG. 4

COHAUSZ & FLORACK
Pat nt- und Rechtsanwälte

COHAUSZ & FLORACK • Postfach 33 02 29 • D-40435 Düsseldorf

Deutsches Patent- und Markenamt

80297 München

Patentanwälte

Dipl.-Ing.	Rudolf Knauf
Dr.-Ing.	Ralph Schippan
Dipl.-Ing.	Andreas Thielmann
Dr. rer. nat.	Ralph Minderop
Dipl.-Ing.	Johannes Simons
Dipl.-Ing.	Hans-Joachim Meyer*
Dipl.-Phys.	Gottfried Schüll
Dr. rer. nat.	Thomas Rox
Dipl.-Ing.	Alexandra Weyres
Dr. rer. nat.	Arwed Burrichter
Dr. rer. nat.	Joachim Brunotte*

Rechtsanwälte

Erik Schäfer
Ulrike Alice Ulrich

Düsseldorf 19. November 2001
Ihr Zeichen
Unser Zeichen WY/ar 010995
Durchwahl +49(0)211-90 49 033

Deutsche Patentanmeldung 101 39 613.9
Silicon Vision AG

A b b u c h u n g s a u f t r a g   Patentsachen

Abbuchungskonto:	04 089 522 00	BLZ	700 800 00
Anmeldegebühr			100,00 DM
Summe der abzubuchenden Gebühren			**100,00 DM**

Gottfried Schüll
Patentanwalt

Kanzlerstrasse 8a
D-40472 Düsseldorf
Telefon +49(0)211-90 49-00
Telefax +49(0)211-90 49-049
*Am Borsigturm 33
D-13507 Berlin
Telefon +49(0)30-43 09 45 60
Telefax +49(0)30-43 09 45 620
mail@cohausz-florack.de
www.cohausz-florack.de